心理科学进展 2021, Vol. 29, No. 4, 677–696
Advances in Psychological Science

https://dx.doi.org/10.3724/SP.J.1042.2021.00677

人际互动中社会学习的计算神经机制*

黎穗卿 陈新玲 翟瑜竹 张怡洁 章植鑫 封春亮

(教育部脑认知与教育科学重点实验室(华南师范大学); 华南师范大学心理学院;
华南师范大学心理应用研究中心; 华南师范大学广东省心理健康与认知科学重点实验室, 广州 510631)

摘 要 人类在社会互动中通过他人的行为对他人特质、意图及特定情境下的社会规范进行学习, 是优化决策、维护积极社会互动的重要条件。近年来, 越来越多的研究通过结合计算模型与神经影像技术对社会学习的认知计算机制及其神经基础进行了深入考察。已有研究发现, 人类的社会学习过程能够较好地被强化学习模型与贝叶斯模型刻画, 主要涉及的认知计算过程包括主观期望、预期误差和不确定性的表征以及信息整合的过程。大脑对这些计算过程的执行主要涉及奖惩加工相关脑区(如腹侧纹状体与腹内侧前额叶)、社会认知加工相关脑区(如背内侧前额叶和颞顶联合区)及认知控制相关脑区(如背外侧前额叶)。需要指出的是, 计算过程与大脑区域之间并不是一一映射的关系, 提示未来研究可借助多变量分析与脑网络分析等技术从系统神经科学的角度来考察大尺度脑网络如何执行不同计算过程。此外, 将来研究应注重生态效度, 利用超扫描技术考察真实互动下的社会学习过程, 并更多地关注内隐社会学习的计算与神经机制。

关键词 社会学习, 计算模型, 神经影像, 强化学习模型, 贝叶斯模型

分类号 B849; C91; B841

1 引言

人类的大部分清醒时间都在与他人互动。在互动中, 人们的社会决策经常依赖于对他人特质、意图以及所处社会情境的认识: 试想你是某重要项目的负责人, 还缺一名同事与你合作, 小李主动请求加入, 你是否会答应让他成为你的合作者? 你开始观察小李工作的效率与质量, 考虑他是否有能力胜任; 你还留意他在工作上是否吃苦耐劳。但由于无法直接观察他人的内在特质, 人们需要在与他人的多次互动中通过他人的行为特点对其能力或道德品质进行判断, 也即在对他人特质进行学习, 以优化决策。当你选择和小李一起开始项目后, 可能需要不断评估对方当前的动机或者行动意愿, 以便适时给予必要的激励。然而, 他人的意图也是无法直接观察的, 相比起较为稳定的能力或道德品质, 意图在动态变化着(Thornton & Tamir, 2017), 为了优化决策, 人们需从他人行为中实时推断他人的所思所想, 即对他人的意图进行学习。最后, 项目成员的合作方式以及最后的成果分配很大程度上受到相应情境下的社会规范的影响。社会规范给人们提供了稳定的心理预期和可靠行为参照, 在社会互动中具有重要作用。然而, 社会规范很大程度上受具体社会情境的制约, 同时社会情境也会影响个体对社会规范的解释, 也即人们信念中固有的规范并非在任何情境下适用, 所以学习特定情境下的社会规范对指导社会决策有着重要的意义。

通过人际互动对他人特质、意图、与特定情境下的社会规范进行学习是人类社会中一种重要的社会学习形式, 构成了人类日常的社会互动中极其重要且复杂的环节。在学习的方式上, 本文关注人类通过与他人直接的互动发生的学习, 同时, 也对在互动情境中的观察学习(observational learning)进行了必要的探讨; 而在学习的目的上(Ruff & Fehr, 2014), 本文讨论的内容主要涉及(1)通过了解他人而更好地惠及自己(也即 for oneself); (2)通过了解他人而更好地在群体中做出合适恰当

收稿日期: 2020-08-10
* 国家自然科学基金(31900757, 32020103008)资助。
通信作者: 封春亮, E-mail: chunliang.feng@m.scnu.edu.cn

的行为(也即 for normative principles)。简括而言，本文主要关注人类在社会互动过程中对他人特质、信念以及社会规范的学习，也即个体为了适应特定的社会互动过程而产生的学习行为。需要指出的是，社会学习也具有其他目的性，例如通过了解他人而更好地惠及他人(也即 for others)。这类社会学习一般称为亲社会学习(prosocial learning)，在人类和动物中均有大量研究，这个研究领域的一个重要发现是扣带前回(anterior cingulate gyrus)特异性地参与了个体的亲社会学习过程(Basile et al., 2020; Hill et al., 2016; Lockwood et al., 2016)。这类研究不在本文的论述范围之内，感兴趣的读者可以参见近年来相关的研究与综述(Apps et al., 2016; Lockwood, Apps, & Chang, 2020; Lockwood, O'Nell, & Apps, 2020)。

社会学习对个体适应复杂的社会环境以及优化社会决策有着重要的意义。首先，社会学习有利于维护自我利益或规避伤害，例如在了解一个人的不诚信品质或欺骗意图后避免与之交易导致上当受骗。其次，个体对他人偏好的学习有利于人际关系的建立与维护。最后，学习特定的社会规范可指引特定情境下的决策，避免触犯规范底线，有助于获得群体的接纳。简而言之，社会学习大大减少了社交中的不确定性，帮助人们优化决策以保障积极的社会互动，是工作与生活顺利进行、维护良好的人际关系以及身心健康的重要条件(Greaves & Farbus, 2006; Loughrey et al., 2020; Seppala et al., 2013)。

由于其重要性与复杂性，社会学习一直是多个研究领域热切关注的主题。近年来，计算模型与神经影像技术越来越多地被应用到社会学习领域的研究中，两种技术的结合体现出独特优势。首先，相比于传统分析只关注个体的外显决策行为，计算模型将社会决策拆解成若干成分，对隐藏的、无法直接测量的社会认知过程予以定量的描述(Suzuki & O'Doherty, 2020; 张银花 等, 2020; 高青林，周媛，2021)，这有利于揭示社会学习所包含的潜在心理计算过程。其次，通过将计算模型里的潜变量与神经影像数据相关联，可揭示执行计算模型所表述的不同计算过程的脑区(Charpentier & O'Doherty, 2018; Cohen et al., 2017; Corrado et al., 2009; Gläscher & O'Doherty, 2010; O'Doherty et al., 2007; Joiner et al., 2017)，这为证实计算模型的合理性提供生物学证据，有利于研究者们了解特定脑区在社会学习过程中的具体参与方式以及执行的计算过程(例如表征主观效用)，使计算模型不只停留在理论假设的层面上。总括而言，计算模型可对社会决策背后的认知计算及其神经机制进行解释，可作为连接神经活动与个体行为表现的桥梁。

简而言之，计算模型为理解社会学习的潜在计算过程与神经机制提供了客观的框架。为了更深入理解社会学习的计算神经机制，本文将对社会学习相关的计算模型及其相关的神经基础进行回顾，以系统地阐述个体在社会互动中对他人特质、意图及社会规范学习的认知计算过程。

2 社会学习的计算机制

2.1 强化学习模型

强化学习的提出源于传统行为主义学说，如条件作用论和强化学习理论，即反复试误后可形成刺激/行为与结果的关联，若某种刺激/行为伴随奖赏结果，个体随后选择该刺激/行为的可能性增加。强化学习模型描绘了自主能动的施动者(agent)为了实现一定的目标，与外界环境(environment)交互时及时根据反馈(reward)性质从而及时调整行动(action)的过程：正性反馈会增加同样行为的几率，而负性的反馈则会削弱后续再次采取同样行为的可能性，如俗话所言，“有借有还，再借不难；有借无还，再借免谈”。在社会情境中，已有的互动反馈(如他人守信/失信行为)会影响后续的社会决策(对他人信任增加/减少)。本小节将介绍在社会学习实证研究领域常见的强化学习模型，从强化学习模型视角介绍社会学习的计算机制。

2.1.1 Rescorla-Wagner 强化学习模型与社会学习

Rescorla-Wagner (RW)强化学习模型(Rescorla-Wagner Reinforcement Learning model) (Rescorla & Wagner, 1972)认为，人类的决策是根据刺激/行为在当前时刻(t)的奖赏期望(expectation, V_t)做出的(Sutton & Barto, 2018)。一个行为带来的奖赏期望越高，其具有的心理效用则更高，因而人们更倾向选择奖赏期望高的决策。若当前的实际结果(outcome, r_t)与预先的期望(V_t)不一致，则计算期望与当前结果(即反馈)之间产生偏差，即预期误差(prediction error, PE_t)。人们将预期误差与原

有的期望加以整合以形成新的期望(V_{t+1}), 新的期望将指引人们的下一次决策。强化学习模型公式可如下表达(公式 1&2):

$$V_{t+1} = V_t + \alpha \cdot PE_t, \quad (1)$$

$$PE_t = r_t - V_t. \quad (2)$$

其中的参数 α 是学习率(learning rate), 取值为 0~1, 代表人们对预期误差赋予的权重, 学习率越接近 1 表示人们在更新期望时越多考虑了近期的结果(Zhang, Lengersdorff et al., 2020)。

RW 强化学习模型最初应用在非社会性学习中。例如, 在经典的强化学习任务里, 两个图形刺激分别对应于获得奖励(果汁)的不同概率, 被试分别更新两个刺激的获奖期望, 进而更多地选择奖赏概率更高的选项(O'Doherty et al., 2004)。又如, 在爱荷华赌博任务中, 被试根据实际收益习得各纸牌的收益期望, 从而选择有利的纸牌(Steingroever et al., 2014)。而在社会互动的情境中, 他人的行为也是有效的反馈信息, 通过比较反馈与期望, 人们不断更新对他人特质或意图的信念。近年来, 大量研究表明个体在社会情境中的学习行为符合强化学习的假设。例如, 在 Chien 等的一项实验里(Chien et al., 2016), 被试和 4 个玩家(A, B, C, D)玩游戏, 每一轮游戏中被试需在两个玩家中选择, 选中的玩家可能给被试一枚金币(1)或不给金币(0)。举例而言, 假设被试对于 4 个玩家的奖赏的初始期望 V_t 均为 0.5。如果某个被试有相对高的学习率(α = 0.8), t 时刻选择了 A 玩家并获得 1 金币, 根据预期误差的计算, 被试对 A 的 PE_t 为 0.5, 经过学习率 α 的缩放后, 对 A 的期望 V_{t+1} 更新为 0.9。经过与 4 个玩家的多次互动, 被试利用反馈逐渐学习到了 4 人各自的慷慨程度, 从而更倾向选择与回报更高的玩家互动。由于社会互动的复杂性, 人们往往采取多样的强化学习模式以更好地优化自己的决策。首先, 人们在不同情境中对结果反馈赋予不同的权重, 也即采取不同的学习率。例如, 人们对于不同吸引力的对象采取不同的学习率, 这说明即使两个人给予自己同等数额的奖励, 两者面孔的吸引力差异导致人们对同样的结果赋予的权重不同(Chien et al., 2016)。此外, 他人面孔可信度对心理效用的作用受其行为的影响, 如果一个人的面孔可信度高并且互惠, 人们认为与其互动有更高的效用, 即在互惠情境下, 面孔可信度增加了人们的效用期望; 然而在背叛情境下, 他人的高面孔可信度与背叛行为产生冲突, 高面孔可信度反而导致更大的效用折扣(Chang et al., 2010)。类似地, 他人为自己提供建议时, 人们会根据建议者的诚实度和信息真实性采取不同学习率来更新信任决策的价值: 当建议者较诚实, 人们对其提供的真实信息比虚假信息赋予更多权重; 相反, 若建议者较不诚实, 人们对其提供的虚假信息结果赋予的权重更多(Bellucci & Park, 2020)。最后, 个体也能根据环境的不确定程度适应性地调整学习率(Franklin & Frank, 2015)。例如, 在信任游戏(Trust game)中, 当委托人返还的金额起伏不定时, 投资人对对方可信程度的学习率会提高, 从而迅速适应社会互动中的高不确定性; 相反, 当委托人返还的金额相对稳定时, 投资人的学习率会下降 (Lamba et al., 2020)。又如, 个体在接受他人建议时, 也会根据对方建议准确性的变化程度来调整学习率, 从而对他人当前的可信程度进行准确的更新(Behrens et al., 2008)。再者, 由于他人特质的不确定性在经过学习后被逐渐消除, 学习率会随着学习的进程逐渐递减(Chien et al., 2016)。

其次, 人们在社会学习中对多个维度同时进行表征与更新。在分配任务里, 人们根据分配者的可支配总额和分配金额, 同时计算对方给予的奖赏(即对方给我多少)与慷慨度(即对方给我的金钱占他可支配金钱的多少)的预期误差, 权衡金钱与慷慨两个维度来更新对对方的期望(Hackel et al., 2015; Hackel & Zaki, 2018); 又如, 在竞争/合作游戏中, 人们根据游戏的输赢结果, 计算对他人能力的预期误差和自己能力的预期误差, 同时更新对自己和他人的能力判断, 并依据此预测与之合作或者竞争可获得的奖赏(Wittmann et al., 2016); 类似地, 人们根据互动结果比较自己已选策略和其他备选策略的优劣性, 并观察他人的决策变化对自己决策策略的影响, 以此追踪他人的决策策略: 若自己原本的最优策略变成次优策略, 则意味着对手改变了策略, 此时人们需及时做出调整, 选出目前对自己最优的策略(Zhu et al., 2012)。再者, 在学习他人的风险偏好时, 人们模拟他人对奖赏的预期, 并依据此预测他人的行为。根据他人的行为和所获得的奖赏结果, 人们可同时计算他人的奖赏预期误差, 以及他人实际行为与自己所估计的出入, 及时更新各选项的奖赏

期望，以准确推断他人接下来的决策(Suzuki et al., 2012)。此外，由于现实社会中的互动常常是重复多次的(我们经常需要和同一个人多次打交道)，因此个体在社会互动中不仅仅需要考虑当前行为带来的即时结果，同时也需考虑行为对将来互动的影响。在公共物品博弈任务(Public Goods Game)中，人们根据组员行为与预期中的偏差，推断组员当前回合合作的可能性，计算个人效用；同时，人们也在计算小组在剩余的游戏回合中合作的长期收益期望，以及自己当前行为对未来合作的影响。也即，人们同时权衡当前收益与长期收益来调整自己当前的合作行为(Park et al., 2019)。

最后，强化学习模型也能够刻画个体对社会规范的学习过程(FeldmanHall et al., 2018; Toyokawa et al., 2019; Zhang & Gläscher, 2020)。例如，在正义游戏(Justice Game)里，第一阶段被试作为接受者面对不公平分配时，更多选择补偿自己而非惩罚分配者；在第二阶段，被试作为第三方替其他接受者做决策，随后查看接受者的实际意愿，但观察到大多接受者更愿意选择惩罚措施，被试在这个过程中通过整合预期误差与原本的信念，学习到他人的规范价值观。被试在第三阶段的任务与第一阶段相同，但由于习得了新的社会规范(也即通过惩罚不公平行为，而非补偿不公平的受害者来维护社会规范)，此时被试更多惩罚不公平的分配者而不是补偿自己(FeldmanHall et al., 2018)。在 Toyokawa 等(2019)的一项群体风险决策实验里，被试每一回合选择三个老虎机中的一个，且被试能看见其他人的决策。在前一阶段，一台老虎机(假设为 A)有较高的赢钱概率，而另外两台(B、C)赢钱概率较低。但在实验的后一阶段，B 成为收益最高的一台老虎机(赢钱概率 B>A>C)。在后一阶段，所有人倾向选择前一阶段最优的选项，即使该选项在后一阶段只是次优的。同时，被试倾向跟随他人的行为，导致群体不能适应环境的变化，“受困于”次优的决策，且这种效应随群体规模增大而增加。对此，Toyokawa 等提出的强化学习模型假设：在群体风险决策中，人们对各选项的奖赏期望进行学习，但人们在群体内的决策实际上是对社会规范(群体某种决策的频率)与私利影响(奖赏期望)的权衡：群体越大，社会影响的权重则越大，这使群体多数人选择的老虎机的效用增加。在三台老虎机收益十分接近，即决策难度较大的情况下，人们对群体决策频率的感知也被放大，此时群体多数人选择的选项的效用也增加(Toyokawa et al., 2019)。

2.1.2 马尔科夫决策过程与社会学习

与传统心理学研究中主要关注环境刺激与个体反应之间的联系不同，马尔科夫决策过程(Markov decision process, MDP)是更完整的强化学习模型。在 MDP 中，处于某一个状态(state)之中的施动者在采取了行动空间中的某个行动 a 后能以一定概率转移到另一个状态中；因此，施动者所处的某个状态 s 与某个特定的行动 a 可共同决定该施动者从原有状态 s 到下一所状态 $\hat{s}$ 的转移，其成功转移的概率称为转移概率 $T(\hat{s},a,s)$ 。

个体在 MDP 中的每一个行动都会影响下一个到达的状态以及获得的奖赏，因此，能使后续累积奖赏期望最高的行动作在当前状态下最佳的决策。个体通过两种学习方法来决定当前状态下的最佳决策：基于模型的(model-based)强化学习和无模型的(model-free)强化学习(Collins & Cockburn, 2020; Dolan & Dayan, 2013)。基于模型的强化学习指的是在明白所有转移概率函数以及奖赏的条件下，计算未来一系列的状态−行为的奖赏。而无模型的强化学习则是用新获得的奖赏与原有的对状态−行动的奖赏预期加权平均，形成新的预期。例如 Daw 等(2011)的二阶段决策任务中，在第一阶段有两个藏语文字选项，被试选择任意选项后以一定概率到达第二阶段中两种状态中的其中一个；第二阶段的两种状态中也各有两个藏语文字选项，分别与不同概率的金钱奖赏相关联，被试在此任务中同时采用基于模型的和无模型的强化学习，也即，既通过求解转移概率函数与奖赏函数来计算当前阶段的各选项的奖赏期望，也根据最新得到的奖励更新对奖励的期望，将两者加权求和。Lockwood, Klein-Flügge 等(2020)将此二阶段决策任务范式改编到社会情境中，发现被试在替自己决策和代替其他人做决策时均同时采取了基于模型和无模型的强化学习，但在替他人做决策时，无模型的学习所占比重更多，而消耗相对多认知资源的基于模型的学习占的比例较少(Lockwood, Klein-Flügge et al., 2020)。

2.2 贝叶斯模型与社会学习

强化学习模型为社会学习过程提供了强有力的解释。按照 RW 模型，社会学习中的内部心理

变量均被计算为确切的数值。然而, 许多研究证实人类能监测并表达自己所感知到的不确定性, 因此 RW 模型所定义的学习方式忽略了人类大脑能够对社会不确定性进行表征(Gershman, 2015)。贝叶斯模型假定人类将感知到的社会不确定性表征在心理变量的后验分布中, 要估计的对象未知且随机, 因此贝叶斯模型能够刻画个体对不确定性的表征。

2.2.1 贝叶斯学习者模型

贝叶斯学习者模型(Bayesian learner model)是描述信息优化过程的一类模型。贝叶斯学习者模型的样式变化丰富, 可应用于不同的社会学习过程中, 但都遵循贝叶斯信息更新方式(公式 3):

$$p(\theta \mid x) = \frac{p(x \mid \theta) p(\theta)}{p(x)} \tag{3}$$

其中, 用 θ 表示待估计的变量(如他人的可信程度), 人们对一切待估计的变量有先验信念 $p(\theta)$, 即自己对某个变量已有的知识; x 表示观测到的数据, 结合先验知识与观测值可求得似然函数 $p(x \mid \theta)$, 似然函数包含了观测样本的所有信息; 根据贝叶斯公式可知, 人们对于待估计变量的后验信念 $p(\theta \mid x)$ 与似然函数和先验分布的乘积成正比(公式 4):

$$p(\theta \mid x) \propto p(x \mid \theta) p(\theta) \tag{4}$$

人们以类似贝叶斯学习者的方式学习他人的特质。比如, 在分配任务里, 被试作为旁观者对分配者的慷慨品质有先验看法, 结合先验信念与实际观察到的分配者行为预测其随后的决策(Stanley, 2016)。又如, 在一项能力学习的任务中, 被试判断专家是否能准确预测资产的涨跌, 人们除了对他人的能力有预先的看法, 也对资产的可预测性有先验信念, 在此基础上结合观察到的他人对资产涨跌的预测以及伴随的结果(预测正确与否)更新对他人能力的认识, 即更新了后验信念, 贝叶斯学习者模型能捕捉到这个信念更新的过程(Boorman et al., 2013)。但该过程受到确认偏差的影响: 人们对专家的能力有较强的积极先验信念, 专家的积极表现(正确预测)比消极表现(错误预测)被赋予更多权重, 影响了信息的整合, 因此, 当一个专家的预测准确率只有 25%时, 人们仍愿意相信其建议(Leong & Zaki, 2018)。此外, 贝叶斯学习者模型还能描述人们学习到他人的权力等级的过程。在 Kumaran 等(2016)的实验里, 被试对公司内的 9 名员工的职位等级进行学习。在被试起初的信念中, 各种等级排序有相等的可能性, 通过观察员工两两比较的结果, 被试不断更新各种职级排序的信念。

人们追踪他人意图的过程也可用贝叶斯学习者模型描述: 比如在群体决策中, 为了让自己做出与其他组员一致的选择, 人们预先估计其他组员对各选项的偏好, 并根据组员们实际的选择, 以贝叶斯的方式更新对他人偏好的看法(Suzuki et al., 2015)。由于互动中的个体之间的意图推断可能是多层次的, 个体可能意识到自己在推断他人的同时, 他人也在推断自己对他的推断, 因此人们在社会互动中的行为可能符合层级贝叶斯学习者模型: 采用低层推断的学习者仅根据观察到的对手的行为更新信念并推断他人未来的行为, 而更高层次的学习者则会推断: “他如何推断我? ”或者“他认为我是如何推断他的? ”因此, 在更高层意图推断者的信念表征中, 并不直接更新他人的行为意图, 而是更新较高层次的“他人对我的推断”, 这些更新通过层次之间的递归计算实现(Devaine et al., 2014; Rusch et al., 2020; Yoshida et al., 2010)。

最后, 贝叶斯模型也能刻画个体对社会规范的整合与适应。当人们处于一个陌生的环境, 环境中大多数人的行为是最直接的信息源, 这些行为被群体中大多数人所认同, 因而可视作该情境下的规范。通常人们按照信念中的规范(也即先验)采取行动, 但如果观察到周围大多数人的决策与自己的决策有偏差, 个体会将这些观察到的规范与人们原本信奉的规范(规范先验)按照贝叶斯方式整合。因此, 规范信念总是朝观察到的规范信息的方向偏移, 从而表现出从众行为(Khalvati, Mirbagheri et al., 2019)。例如, 在跨期选择或者风险决策任务中, 被试从其他人的行为中观察到多数人的选择偏好, 并将自己的偏好(先验信念)与多数人的行为偏好(社会规范)以贝叶斯信念更新的方式进行整合, 从而让自身的行为与群体规范相适应(Garvert et al., 2015; Reiter et al., 2019; Suzuki et al., 2016)。又如, 风险规避的个体按照自身原本的信念倾向在风险决策中选择保守的选项, 但当观察到其他人均在追求风险, 被试则将“风险寻求”视作当下情境的一种规范, 因而改变原本的风险偏好来适应当前的规范(Suzuki et al.,

2016)。相比起成年人，青少年的风险偏好更易受到同龄人行为的影响，具体表现为：当同伴均在寻求风险时，青少年更多将这种规范整合到信念中去，信念改变的幅度更大(Reiter et al., 2019)。在最后通牒(Ultimatun game)游戏里，受公平信念影响，应对者在游戏开始阶段均期望获得 50%的分配(先验信念)，因而较多地拒绝不均等分配。但随着游戏进行，人们以贝叶斯观察者的方式及时从他人的行为中习得新的公平规范，并调整预期。例如，当应对者发现多数分配者给出的分配均不符合公平原则，如分配均值为 20%，应对者会适应性地降低期望，更多地接受不公平的分配。相反，若分配的均值在 60%，应对者的期望提升，对于后续相对公平的分配(40%)，应对者也更多地拒绝(Gu et al., 2015; Xiang et al., 2013)。值得注意的是，与贝叶斯模型的假设一致，个体将群体观点整合到自己信念的过程受到自身信念以及社会信息可靠程度的影响(de Martino et al., 2017; Park et al., 2017; Toelch et al., 2014)。一方面，规范先验信念较弱的个体更易受到多数人行为的影响而更多更新信念；另一方面，群体人数越多时，群体行为越是可靠的规范证据，先验信念更可能被改变。例如，被试要求在阅读一段犯罪描述后需决策给罪犯判处多少年的刑罚。首次评分确定性越低的被试，越容易受到评审团其余人的判决的影响而改变决策；当评审团人数增加，被试也更容易改变决策，使自己的判决与评审团的平均分接近(Park et al., 2017)。

2.2.2 分层高斯过滤器

社会情境中最直接的不确定性来源于我们无法得知他人接下来的行为，而这种不确定性受到许多他人内部状态(hidden states)的影响。比如，当我们对街头的流浪者予以金钱施舍(x_1)，该行为可能出于我们想要帮助他人的倾向(x_2)的影响，这种倾向进一步受到利他动机的波动(x_3)的影响。又如，梁山好汉盗取他人金钱的行为(x_1)出于获得金钱的目的(x_2)，但该目的是由劫富济贫的意图(x_3)所产生的，这种意图可能由助人动机(x_4)所导致。由这些例子可见，直观的行为依赖于复杂的、不可见的内部状态，而每一层次的内部状态又依赖于更高层的状态(最高层次状态除外)，这种低层次状态对高层次状态的依赖关系在分层高斯过滤器(Hierarchical Gaussian Filter, HGF)模型中称之为耦合(coupling)。HGF 假设，人们观察到的他人的行为结果(x_1)均取决于他人更高层次的内部状态($x_2, x_3, x_4 \ldots\ldots x_n$)。这些内部状态在时间上以高斯随机漫步(Gaussian random walks)的形式变化：除了最高层的状态 x_n 以外(x_n 不依赖于其他层，$x_n^k \sim N(x_n^{k-1}, \theta)$)，人们在某时刻($k$)对他人行为/状态的表征 x_i^k 服从高斯分布，该分布的均值为上一时刻(k–1)对该行为/状态的表征 x_i^{k-1}，分布的方差是关于更高层次状态的函数 $f_i(x_{i+1})$，即 $x_i^k \sim N(x_i^{k-1}, f_i(x_{i+1}))$，该函数反映了上一层状态对下一层状态的稳定性的制约，制约关系由若干个参数决定(C. Mathys et al., 2011; C. D. Mathys et al., 2014) (公式 5)：

$$f_i(x_{i+1}) \stackrel{\text{def}}{=} \exp(\kappa_i x_{i+1} + \omega_i) \tag{5}$$

由函数表达式可知，ω_i 是第 i 层状态 x_i 独立于更高层状态的变量，表示 x_i 的稳定性，ω_i 越大则表示 x_i 的不确定性越强；κ_i 是上一层状态 x_{i+1} 的系数，因此决定了 x_i 与上层状态 x_{i+1} 的耦合紧密程度，κ_i 越大，x_i 与 x_{i+1} 的依赖越紧密。最高层状态 x_n 不依赖于其它层并对底层状态起决定作用，其波动性受其分布的方差 θ 制约，θ 越大则表示最高层状态的波动性越强。

总的来说，HGF 有三大优点。首先，HGF 将人们感知的状态以概率分布的形式表征，捕捉表征的不确定性。其次，HGF 强调了心理计算的分层性。HGF 在贝叶斯学习者的基础上加入影响下层心理变量的上层因素，对人们感知的不确定性的来源进行解释。最后，HGF 刻画的学习过程反映了人们在不确定性强度不同的情境下动态的学习速率(C. Mathys et al., 2011; C. D. Mathys et al., 2014)。由预期误差驱动的学习发生在各层次中，各层的学习率是当前感知到的状态稳定性与状态的先验稳定性的比值。因此，对状态先验的表征越不稳定(即状态先验的方差越大)，或当前感知到的状态稳定性越强，人们对预期误差赋予更多的权重。由于各层次的状态在时间上以高斯随机漫步的形式变化，学习率时刻反映着表征中的不确定性与环境中的不确定性。

HGF 早期主要是对知觉任务进行建模(Lawson et al., 2017; Powers et al., 2017)，但近年来越来越多的研究者开始将 HGF 模型用于对社会学习行为的拟合。Siegel 等(2018)用 HGF 模型捕捉到人们对他人道德品质的学习。被试观察他

人是否会因为金钱利益而对别人实施电击伤害，两层的 HGF 模型刻画了被试的学习过程：低层是他人的行为结果(x_1)，取决于其高层状态——道德品质的好坏(x_2)，因此人们通过学习其道德品质及稳定性，对其接下来的行为做出准确的预测。x_2 不依赖于其它层，其不确定性由其服从正态分布的方差 exp (ω)反映。模型拟合的结果表明人们对坏人的道德品质表征有更大的方差，因此是更不稳定的(Siegel et al., 2018)。在另一项建议采纳的实验里，他人提供的建议的真实性(x_1)受到其助人动机(x_2)的影响，助人动机进一步受到动机波动(x_3)的制约，最高层 x_3 的不确定性独立于其它层，方差为 θ。被试根据他人意图的变化来推断其建议的真实性，以便优化他们对建议有效性的预测：当他人持续提供误导性信息，人们极少听从对方的建议，在模型拟合结果中表现为对对方的 x_3 方差 θ 的表征更大，即对其助人动机的感知十分不稳定(Diaconescu et al., 2014; Diaconescu et al., 2017)。类似地，在风险决策任务里，他人目光注视的方向给人们的决策提供暗示，暗示的正确性(x_1)受线索提供者提供真实线索的动机(x_2)的影响，这种动机又进一步受到动机不稳定性(x_3)的制约(Henco et al., 2020)。值得注意的是，对他人稳定性的感知受社会偏好的影响：社会偏好强的人给社会信息(也即他人目光注视方向)的稳定性赋予更多权重，将社会信息的可靠程度放大，因此更可能采用他人提供的线索而不顾自己掌握的信息(Diaconescu et al., 2020; Henco et al., 2020)。

2.2.3　卡尔曼滤波器

可捕捉学习过程中的不确定性的模型还有卡尔曼滤波器(Kalman filter)也。在卡尔曼滤波器中，对奖赏的预期是经验中的刺激−奖赏联结与联结权重的总和，因此预期不仅基于刚刚发生的结果，而取决于所有已获取的观测。刺激−奖赏联结的权重在每一次获取观测值后都得到更新，这种更新由观测值与预期之间的误差所驱动，更新速率与预期信念的不确定性有关。因此，预期误差越大，或预期不确定性越大的时候，预期改变得越快。目前，卡尔曼滤波器主要在非社会学习的相关研究中得到应用和发展(Yu & Dayan, 2003; Dayan et al., 2000; Morris et al., 2017; Piray & Daw, 2020)，但尚未被应用于有关社会学习与社会决策的研究中，研究者们在今后的研究中可尝试将卡尔曼滤波器与其他数学模型进行比较，挖掘更多解释社会学习计算机制的可能性。

2.3　强化学习模型与贝叶斯模型的结合——部分可观察的马尔科夫决策过程模型

部分可观察的马尔科夫决策过程(partially observable Markov decision process, POMDP)的理念源于 MDP，因此具备强化学习的属性。然而在真实的社会情境中，人们不能准确掌握环境的真实的状态(如他人的道德品质)，仅能通过感官获得一部分能反映状态的观测值(如外显的行为)，此时人们利用观测值形成对真实状态的概率估计，MDP 则变成部分可观察的 MDP (partially observable MDP, POMDP)。除了 MDP 包含的元素，POMDP 还包括了人们获得的观测值 Z，及在状态先验条件下的观测值似然函数 O。在 POMDP 中，人们根据状态的后验信念，也称信念状态(belief state)做出决策；而对信念状态的更新则基于遵循贝叶斯信息更新方式，因此 POMDP 也具备贝叶斯的属性。Khalvati, Park 等(2019)发现 POMDP 比强化学习模型能更好拟合人们在小组合作任务中的决策过程：人们根据历史互动结果形成信念状态(对组员合作程度的概率估计)，预测接下来小组中可能选择合作的人数，进而选择最有利的行为。此外，在 POMDP 中，每一次决策后发生的信念状态的转移(后验信念的更新)都是下一次决策的起点；因此，基于更新得到的对组员合作性的看法，人们权衡个人当前的效用与小组在剩余回合的长期收益期望，采取长远的眼光计划接下来的决策，而不仅仅关注当下回合的决策(Khalvati, Park et al., 2019)。

交互式部分可观察的马尔科夫决策过程(Interactive POMDP, IPOMDP)在 POMDP 的基础上更多地考虑了社会互动过程中的真实情况。尽管 POMDP 刻画了个体利用互动中的反馈更新信念状态的过程，但它忽略了一个事实：互动的对象可能也在以同样的方式形成对自己的信念状态，也即人们会意识到他人也在对自己进行意图的推断，或者更高阶的推断：他认为我如何对他进行推断？社会情境的真实状态受多个施动个体行为的影响，因此对个体行为的预测是对状态估计的必要条件(Gmytrasiewicz & Doshi, 2005)。IPOMDP 与 POMDP 最大的区别在于 IPOMDP 模型里对状态的估计还包括了对他人模型的估计，POMDP 中

的状态空间在 IPOMDP 中变成与其他施动者模型有关的交互式状态空间(Hula et al., 2015; Hula et al., 2018; Kumar et al., 2019; Xiang et al., 2012)。例如，在多轮信任游戏里，投资者(i)通过从互动历史中学习委托人(j)的行为特点，计算当前回合下，若自己选择了行动空间(A_i)中某个决策(a_i)后，委托人做出他的行为空间(A_j)中的某种决策(a_j)的可能性及其效用，提前估计在未来两个回合中的收益，以优化自己的决策(Xiang et al., 2012)。此外，除了实时推断他人的互惠偏好和策略，提前规划决策外(Hula et al., 2015)，人们还从他人的行为中推断他人当前的情绪，当发现自己的决策引起他人愤怒时便及时调整策略以修复合作关系(Hula et al., 2018)。简而言之，IPOMDP 刻画的是多个采用 IPOMDP 的个体多次互动的过程，即他人的意图模型也是决策过程需考虑的因素。

综上，强化学习模型与贝叶斯模型各有其特点：强化学习模型计算简便，并通过学习率参数刻画不同个体之间的学习差异，然而这种计算方式未能有效表征人们对社会情境中的不确定性的感知；贝叶斯学习模型所描述的学习是一种基于概率论的、信息优化过程，并采用概率分布来描述心理变量，更能反映人们对不确定性的表征。需要指出的是，以往许多研究发现强化学习和贝叶斯的算法也适用于对非社会学习的描述(Devaine et al., 2014; Lockwood et al., 2018; Stanley, 2016)，这说明社会学习与非社会学习的认知计算过程有一定程度的相似性。但需注意的是，两者仍存在显著的差别：与非社会学习相比，社会学习通常涉及心理理论等社会认知过程，也即人们在社会互动中不仅追踪互动产生的奖赏或惩罚，还考虑了个体间行为的相互影响，因此，相比于标准的强化学习或贝叶斯模型，考虑了社会影响因素的模型往往能更好地对社会学习进行描述(Hampton et al., 2008; Hill et al., 2017; Zhu et al., 2012; Park et al., 2019; Zhang & Gläscher, 2020)。

2.4 计算模型的选择

需要指出的是，强化学习模型与贝叶斯模型以及它们的变式适用于描述个体在各种环境中的学习行为(Pulcu & Browning, 2019; Soltani & Izquierdo, 2019)；而目前也尚不清楚哪种模型更真实地反映了学习的内在机制(Pulcu & Browning, 2019)。因此，在实际的使用中，对模型的选择主要借助模型比较指标，这些指标反映了对模型复杂度以及模型数据拟合度的权衡(Wilson & Collins, 2019)。例如，通过模型比较 Siegel 等发现 HGF 模型优于 RW 强化学习模型(Siegel et al., 2018)。然而，借助常用的模型比较指标只能得出不同备选模型的相对优劣，而不能评估模型表现的绝对优异程度(Palminteri et al., 2017)。这个问题可以通过参数恢复(parameter recovery)与后验预测检验(posterior predictive check)的方式解决，这种方法已经在越来越多的研究中被采纳(Khalvati, Park et al., 2019; Palminteri et al., 2015)。最后，也有研究者借鉴“开放科学(open science)”领域的经验对模型建构过程的可靠性提出了建议。例如预注册自己的研究会涉及的模型，使用的模型比较指标，公开建模的过程，等等(Lee et al., 2019)。由于建模过程本身是比较多样和灵活的，这种措施能提高其中的透明度，在一定程度上避免计算模型被滥用和误用。

2.5 计算模型的实现

计算建模涉及到一定的数理和编程知识，对初学者来说有一定难度，但可借助现有的工具包和软件。其中 hBayesDM toolbox 是基于 R 和 Python 的工具包，通过它们编译和调用 Stan 模型，使用马尔科夫链蒙特卡洛方法(Markov Chain Monte Carlo)对模型的参数分布进行采样，并使用层级贝叶斯方法(hierarchical Bayesian approach)同时对群体和个体水平的参数进行采样(Ahn et al., 2017)。该工具包收录了大量经典决策和学习模型，因此，如果使用了相同的范式进行研究，可直接方便地调用相关的函数对自己的数据进行建模分析。此外，HGF toolbox 是基于 Matlab 的工具包(Mathys et al., 2011; Mathys et al., 2014)，收录了常用的强化学习模型(如 RW 模型)以及贝叶斯模型(如 HGF 模型)，使用最大后验(maximum-a-posteriori)的方法进行参数估计。再者，VBA toolbox 也是基于 Matlab 的工具包(Daunizeau et al., 2014)，可用于对神经生物学以及行为数据进行建模，并使用变分贝叶斯方法(variational Bayesian approach)进行参数估计。VBA toolbox 中收录了与学习以及决策相关的强化学习与贝叶斯模型，它的核心思想是将所有模型拆分为状态演化方程(evolution equation)以及状态–观测映射方程(observation equation)，因此不同的模型能以统一的方式进行

调用。最后, CBM toolbox 是基于 Matlab 的工具包, 与主流模型比较方法假设某个最优模型适用于所有被试不同, CBM 最大的特点是允许不同被试可能对应不同的最优模型(Piray et al., 2019)。这个方法将模型比较与参数估计统一到一个框架下, 因此能够得到更为准确的结果。虽然这种方法目前使用得不多, 但很可能是未来研究的重要方向。但是, 这种方法要求使用者知道自己的模型空间是什么(总共有哪些模型), 并将它们编写出来, 因此对编程的要求较高。

总的来说, 现有的工具箱为研究者使用计算建模分析提供了一定程度的便利, 研究者可以至少在三个层面上使用这些工具箱。首先, 研究者可能使用了工具箱里已经收录的实验范式和对应的模型, 这时可直接调用包内的对应函数。其次, 研究者可能需要根据自己的任务增加工具箱中没有的模型, 这时工具箱里的模型函数的源代码依然可以作为编程语法和逻辑上的模板, 帮助研究者在现有模型基础上改编自己所需的模型。最后, 大多工具箱都提供了模型比较和参数估计的模块, 所以即使研究者使用的是完全不同的模型, 依然可以借用工具箱中的这些模块, 来减少编程上的代价。

3　社会学习的神经机制

近年来, 越来越多的研究将计算模型与脑成像技术结合, 以此探究大脑对社会学习中不同认知计算过程的执行方式。这种基于模型的脑机制研究一方面有助于理解大脑各区域在社会学习中的具体作用, 另一方面可以为计算模型提供生物学层面的支撑。本小节将对有关社会学习的计算神经影像研究进行回顾, 从不同类型的计算模型的角度分别归纳社会学习的脑机制。

3.1　强化学习的神经机制

强化学习模型假设决策基于心理效用, 因此在决策过程中人们需要对各决策选项的效用进行计算, 心理效用的表征发生在腹内侧前额叶(ventromedial prefrontal cortex)、眶额叶(orbital frontal cortex)和腹侧纹状体(ventral striatum)等负责奖赏与价值编码相关的脑区中(Behrens et al., 2008; Bellucci et al., 2019; Chien et al., 2016; Hackel et al., 2015; Zhang & Gläscher, 2020)。这些脑区表征的价值包括两方面。首先, 这些脑区表征了社会互动中获得的非社会性效用。例如, 腹内侧前额叶和腹侧纹状体表征了对互动时金钱收益的计算(Behrens et al., 2008; Park et al., 2019; Zhang & Gläscher, 2020): 人们在建议采纳任务中选择更优选项时伴随着更强的腹侧纹状体活动(Bellucci et al., 2019); 获得他人给予的金钱时, 腹内侧前额叶和腹侧纹状体活动会增强(Chien et al., 2016)。以往非社会学习的研究中也发现这些脑区对物质奖赏的编码作用, 例如被试在刺激−奖赏关联学习任务中获得奖励时腹内侧前额叶激活增加(Blair et al., 2006; Jocham et al., 2011)。此外, 这些价值编码相关的脑区(如腹内侧前额叶)还表征了社会学习中的社会性效用, 如良好的特质(如诚实、慷慨)、意图(助人动机)以及社会规范等本身固有的心理价值。例如, 面孔吸引力对于人们来说是种奖赏, 人们与高面孔吸引力者(Chien et al., 2016)互动时腹内侧前额叶有更强的激活。类似地, 与诚实的人(Bellucci et al., 2019)互动时腹内侧前额叶激活更强。在选择互动的伙伴的时候, 人们对他人的慷慨程度进行比较, 慷慨的差异表征在腹内侧前额叶中(Hackel et al., 2015)。最后, 腹内侧前额叶损伤的病人不能很好地计算不同公平程度的分配所具有的效用差异, 因而更多接受不公平的分配, 让自己处于不利地位(Gu et al., 2015)。

但在社会情境下的奖赏不如非社会情境下那样可直接观测, 社会奖赏的计算经常基于对他人特质或意图的信念, 因此需要借助外显行为进行推断方可获得。所以社会学习的效用的表征时常涉及社会认知加工相关的脑区, 如后扣带皮层(posterior cingulate cortex)以及颞上沟(superior temporal sulcus)、颞顶联合区(temporo-parietal junction)、背内侧前额叶(dorsomedial prefrontal cortex)等脑区的参与(Bellucci et al., 2019; Ligneul et al., 2016; Park et al., 2019; Wittmann et al., 2016)。例如, 在竞争游戏里, 人们推断他人相较于自己的能力支配地位, 以便选出更易击败的对手, 背内侧前额叶表征了对他人能力的预期(Ligneul et al., 2016); 而经颅直流电对背内侧前额叶活动的抑制破坏了这种表征, 从而使人们更多选择比自己地位高的人作为对手。与之类似, Wittmann 等(2016)发现背内侧前额叶的信号追踪他人的能力水平。又如, 在合作任务里, 群体利益

依赖于各群体成员选择贡献/搭便车的可能性，因此人们在计算群体长期利益时需对其他组员的行为动机进行推断，相应地伴随着颞顶联合区更强的活动(Park et al., 2019)。同理，在做出信任决策之前对效用的计算依赖于对他人诚实特质或意图的推断，涉及心理理论的加工，对他人特质或意图的推测主要编码在后扣带皮层中；他人的诚实行为增加了颞顶联合区与腹内侧前额叶的功能连接，而两者的连接强度可预测被试对他人的信任程度(Bellucci et al., 2019)。

强化学习发生的关键在于预期误差的驱动作用。预期误差的计算需要将预期的心理效用与实际观察结果进行比较，主要涉及腹内侧前额叶、眶额叶、壳核(putamen)以及腹侧纹状体等奖赏加工相关脑区(Behrens et al., 2008; Hackel et al., 2015; Zhang & Gläscher, 2020; Zhu et al., 2012)以及前扣带皮层(anterior cingulate cortex)和前侧脑岛(anterior insula)等冲突与错误加工相关脑区(Zhang & Gläscher, 2020; Zhu et al., 2012)。例如，当他人给与自己的金钱超出自己的预期时(Chien et al., 2016; Hackel et al., 2015)，或在任务里赢得的钱比预期中更多时(Behrens et al., 2008; Zhang & Gläscher, 2020; Zhu et al., 2012)，腹侧纹状体、壳核和前扣带皮层的活动增强。类似地，当自己或慈善机构获得的钱比预期要多，腹侧纹状体的活动增强(Kuss et al., 2013)。在 Jones 等(2011)的一项实验里，被试通过写便签的方式表示对他人(生活、爱好等)的兴趣，当被试在不期待得到对方反馈时却收到对方写给自己的便签时，前扣带皮层、前侧脑岛、腹侧纹状体和眶额叶的活动增强。最后，脑岛损伤的病人对预期误差过度敏感，而不能良好地适应社会规范的变化(Gu et al., 2015)。

监控社会情境中预期效用与实际结果的冲突还依赖于社会认知能力，即对他人特质、意图的推断，该过程还涉及到颞上沟、颞顶联合区、背内侧前额叶、楔前叶(precuneus)以及颞中回(middle temporal gyrus)等脑区(Behrens et al., 2008; Chien et al., 2016; Hackel et al., 2015; Wittmann et al., 2016; Zhang & Gläscher, 2020; Zhu et al., 2012)。例如，对分配者慷慨程度的预期误差的计算需要从其行为推断其品质，这一认知过程引起后扣带皮层、楔前叶和颞顶联合区更强的活动，但这些脑区并没有参与到金钱预期误差的计算中(Hackel et al., 2015)。类似地，若他人提供信息的真实性与预期不一致，颞中回、背内侧前额叶、颞上沟和颞顶联合区活动增强(Behrens et al., 2008)；他人的策略发生变化时，背内侧前额叶参与表征对对手信念判断的错误(Zhu et al., 2012)。在竞争或合作游戏中，背内侧前额叶追踪他人的能力水平，并计算对能力估计的误差(Ligneul et al., 2016; Wittmann et al., 2016)。在群体决策里，当人们观察到自己的决策与多数人的决策不一致时，颞顶联合区和壳核的功能连接增强，并且它们的连接强度可以预测人们随后的策略变化(Zhang & Gläscher, 2020)。

除了奖惩加工相关脑区与心理理论加工相关脑区，也偶有研究发现了其他脑区参与到了主观效用与预期误差的计算。例如，在推断他人特质时，背外侧前额叶(dorsolateral prefrontal cortex)也参与了对他人诚实特质表征(Bellucci et al., 2019)，而腹外侧前额叶(ventrolateral prefrontal cortex)编码了人们对他人慷慨特质预测的误差(Hackel et al., 2015)。对他人策略的预期误差还会涉及枕叶皮层(occipital cortex)和小脑(cerebellum)的参与(Zhu et al., 2012)。

3.2 贝叶斯信念更新的神经机制

贝叶斯框架下的社会学习中，人们与他人互动之前便已对他人特质与意图或社会规范有先验信念。例如，在最后通牒游戏开始前，应对者预期分配者给出的分配约为总额的 50% (Xiang et al., 2013)。先验信念的形成主要依赖于对他人的所思所想进行的推断，以往研究发现该认知过程的执行主要涉及心理理论相关和奖赏加工相关的脑区(Boorman et al., 2013; Kumaran et al., 2016; Suzuki et al., 2015)。比如，人们对他人提示信息(注视方向)的正确性表征在颞下回(inferior temporal gyrus)、苍白球(pallidum)和壳核(Henco et al., 2020)；对他人能力水平的估计表征在前扣带皮层、后扣带皮层和背内侧前额叶(Boorman et al., 2013)。小组任务里，被试选择自己喜爱的物品时，主动回忆该物品在之前受其他组员的喜爱程度，同时推断当前各组员对不同物品的喜爱，这个计算过程主要由颞上沟和颞顶联合区负责(Suzuki et al., 2015)。在权力等级的学习中，杏仁核(amygdala)、海马体(hippocampus)和腹内侧前额叶参与了对他人权力的表征(Kumaran et al., 2016)。

通过比较模型中信念的均值与实际反馈结果的差异，人们的先验信念的偏差可被量化，类似于强化学习中对预期误差的计算。实际上，研究者们也发现在贝叶斯框架下的社会学习中，大脑的确计算了信念分布的均值与实际结果之间的差异。一方面，对他人行为或策略的预期表征在与奖赏编码相关的脑区以及感知冲突、负性情绪的脑区中(Diaconescu et al., 2017; Garvert et al., 2015; Henco et al., 2020; Xiang et al., 2013; Xiang et al., 2012)。例如，委托人的互惠金额与信念之间的差异、以及自己的投资与他人对自己的投资预期的差异均表征在腹侧纹状体的活动中(Xiang et al., 2012)。又如，对他人诚实度和博弈策略的信念误差增加了前扣带皮层与前侧脑岛的活动(Diaconescu et al., 2017)。他人行为对公平的违背则反映在腹侧纹状体、腹内侧前额叶、眶额叶和前侧脑岛的活动中(Xiang et al., 2013)。另一方面，信念偏差计算过程也涉及对他人潜在心理状态的推断，因此经常涉及心理理论加工相关的区域，包括颞上沟、颞顶联合区、楔前叶和背内侧前额叶(Boorman et al., 2013; Diaconescu et al., 2017; Stanley, 2016; Xiang et al., 2012)。例如，人们对专家专业水平的信念和专家实际能力的比较发生在颞上沟、颞顶联合区、背内侧前额叶、颞中回、楔前叶和运动前区皮层(promotor cortex) (Boorman et al., 2013)。类似地，比较自己投资与他人对自己的投资期望的差异涉及颞顶联合区的参与(Xiang et al., 2012)。由于贝叶斯模型假设信念的表征服从概率分布而不是一个确切的数值，因此在社会学习过程中，人们还对信念的不确定性进行表征。目前发现的表征不确定性的脑区主要包括奖赏/情感加工相关区域、心理理论相关脑区以及认知控制相关区域(Diaconescu et al., 2020; Henco et al., 2020; Hétu et al., 2017; Xiang et al., 2013; Yoshida et al., 2010)。例如，在建议采纳任务中，决策基于对自己掌握的非社会信息(赢钱概率)与社会信息(他人建议)的可靠性的权衡。Behrens 等(2008)发现人们对社会性与非社会性的不确定性表征在前扣带皮层的不同位置，其中前扣带沟负责感知金钱奖赏的波动，而前扣带回表征了他人意图的不稳定性。Diaconescu 等(2020)则发现，人们对非社会信息和社会信息的可靠性赋予不同的权重，当人们更依赖自我掌握的信息的可靠性做决策，腹侧被盖区/黑质(ventral tegmental area/ substantia nigra)和背外侧前额叶有更强的激活；相反，当人们给社会信息的可靠性赋予更多权重时，腹内侧前额叶和颞顶联合区有更强的活动；此外，与意图稳定的人互动时杏仁核、颞上沟和腹内侧前额叶有更强的激活。在类似的任务中，Henco 等(Henco et al., 2020)发现社会信息的不确定性主要表征在颞上沟，而非社会信息的不确定性主要表征在前侧脑岛与中央盖(rolandic operculum)中。又如，在合作游戏里，当对手(程序)的行为策略改变，人们对他人策略推断的不确定性增加，伴随着内侧前额叶与后扣带皮层有更强的激活，说明内侧前额叶和后扣带皮层表征了信念的不确定性(Yoshida et al., 2010)。

另外，大脑还表征当前环境下社会规范的不确定性：人们规范信念的分布方差的误差主要表征在前侧脑岛、腹侧纹状体和腹侧被盖区/黑质(Hétu et al., 2017; Xiang et al., 2013)。由于社会规范信息与人们先验信念以贝叶斯信息更新方式整合，在此框架下，个体需对社会信息与个人信息的可靠性(与不确定性互为倒数)进行评估。研究发现信息的可靠性主要表征在腹内侧前额叶、腹侧纹状体、颞顶联合区以及楔前叶等脑区(Burke et al., 2010; Campbell-Meiklejohn et al., 2017; de Martino et al., 2017; Li et al., 2019; Meshi et al., 2012; Park et al., 2017) 。例如，腹内侧前额叶和背内侧前额叶表征了人们对某件商品评分的肯定程度(de Martino et al., 2017)。类似地，在猜瓮迷游戏里，腹内侧前额叶表征了他人对多数石头颜色判断的肯定程度(Campbell-Meiklejohn et al., 2017)。观看他人购买股票行为时，购买人数越多意味着社会信息的可靠性越高，腹侧纹状体有更强的激活；人类的行为比猩猩的购买行为更具参考价值，在观看人类购买股票行为时腹侧纹状体活动更强(Burke et al., 2010)。

最后，人们根据观测到的信息的可信度，将之与先验信念进行整合，从而更新自己的信念并据此做出合适的决定。在贝叶斯模型中，信念更新的程度可以用 Kullback-Leibler divergence (KL 散度)衡量。KL 散度是两个概率分布之间差异的度量，因而可用于描述后验信念与先验信念之间的差异，即信念更新的程度。信念更新的程度主要与认知控制加工相关区域的活动呈正相关，包

括背外侧前额叶、顶下小叶(inferior parietal lobule)以及前扣带皮层(de Martino et al., 2017; Park et al., 2017; Suzuki et al., 2016)。例如，在规模越大的群体中，被试越容易将群体内多数成员的决策视作有效信息从而做出更大的从众行为，这种行为上的变化能够被背外侧前额叶和前扣带皮层功能连接强度预测(Park et al., 2017)。又如，在小组决策过程中，表征个人物品喜好的腹内侧前额叶和表征组员喜好的颞上沟/颞顶联合区与前扣带皮层功能连接增加，而前扣带皮层的激活强度与被试的物品选择概率相关，说明前扣带皮层整合了被试的个人喜好以及对组员喜好的估计，并指导被试的决策(Suzuki et al., 2012)。此外，研究者也发现奖赏系统参与到信息整合的过程。例如，对自己相对于他人权力等级的信念更新的程度与眶额叶、海马体、前侧脑岛及面孔加工相关脑区梭状回面孔区(fusiform face area)的激活相关(Kumaran et al., 2016)。又如，腹侧纹状体的活动可预测人们在购买股票时受群众影响的程度(Burke et al., 2010)。类似地，在跨期选择任务中，被试自己的延迟满足偏好与他人的差异计算在腹侧纹状体中，并且腹侧纹状体对差异的表征强度可预测被试延迟满足偏好的改变程度(Garvert et al., 2015)。又如，对自己与他人权力等级的信念更新的程度与海马体、眶额叶、前侧脑岛及梭状回面孔区(fusiform face area)的激活相关(Kumaran et al., 2016)。

总而言之，在强化学习框架下，社会学习主要包括对主观期望与预期误差的计算。与非社会学习的研究发现一致，以上两个计算过程涉及奖惩加工相关脑区的参与；但在社会性情境中，由于需要通过行为对他人的特质与意图进行推断，社会认知加工相关脑区也参与到主观期望和预期误差的计算中。而在贝叶斯模型下，人们表征对他人的先验信念，并感知到实际结果与信念的偏差，其表征分别与强化学习中主观期望、预期误差的表征相类似。此外，在贝叶斯模型框架下的社会学习还涉及对不确定性的表征以及信息整合，除奖惩加工和社会认知加工脑区外，还涉及认知控制相关脑区。这些研究发现提示社会学习与非社会学习的神经机制有重叠部分，主要表现在奖惩加工相关区域在学习中的价值表征作用；然而，重要的是，社会学习相关计算过程的执行也涉及社会加工的特异性脑区，例如社会认知相关脑区对学习信号的计算(O'Doherty et al., 2017; Ruff & Fehr, 2014)。

4 总结

社会学习，即在反复进行的社会互动中对他人的特质、意图以及社会规范的学习，能有效地帮助人们降低社会互动中的不确定性并优化决策。计算模型、行为范式以及神经影像技术的结合可将社会学习背后复杂的认知计算过程量化并揭示执行这些计算过程的大脑区域。因此，近年来基于模型的脑成像研究在很大程度上加深了我们对社会学习计算神经机制的理解。

本文对近年来有关计算建模以及计算神经科学相关的实证研究进行了回顾，主要关注了现有文献中主要使用的两大类计算模型—强化学习模型和贝叶斯模型—如何刻画个体的社会学习过程及其神经基础。强化学习模型假设社会学习由实际互动结果与期望之间的误差驱动，人们对预期误差赋予一定的权重(学习率)后与原有期望整合，而学习率捕捉到社会学习过程的个体差异。贝叶斯模型则假设社会学习的过程符合贝叶斯信息更新的方式：人们将观察到的现象与对他人或规范的先验信念以贝叶斯的方式进行整合，整合形成后验信念将指导后续的社会决策。

与强化学习以及贝叶斯模型密切相关的潜在计算过程主要包括对主观期望以及预期误差的表征、对不确定性的表征以及对信息的整合过程。这些核心的计算过程涉及的脑区包括奖惩加工、心理理论以及认知控制相关的脑区(图 1)。具体而言，主观期望主要表征在奖赏加工相关脑区与心理理论加工相关脑区，对预期误差表征的表征主要发生在奖赏加工相关脑区、情感加工相关脑区以及心理理论加工相关脑区。主观期望和预期误差的计算在强化学习模型与贝叶斯模型中均有涉及，但贝叶斯模型下的社会学习还涉及对不确定性的表征与信息整合，两者均涉及奖赏加工相关脑区、情感加工相关脑区、心理理论加工相关脑区和认知控制相关脑区的参与。现有的证据提示两个主要的结论：首先，在算法层面，虽然强化学习与贝叶斯模型同时适用于社会和非社会学习；但在执行层面，可以发现社会学习涉及的认知计算过程(如对预期误差的表征)不仅仅依赖于非社会学习中的奖惩加工区域，同时也涉及与社会认

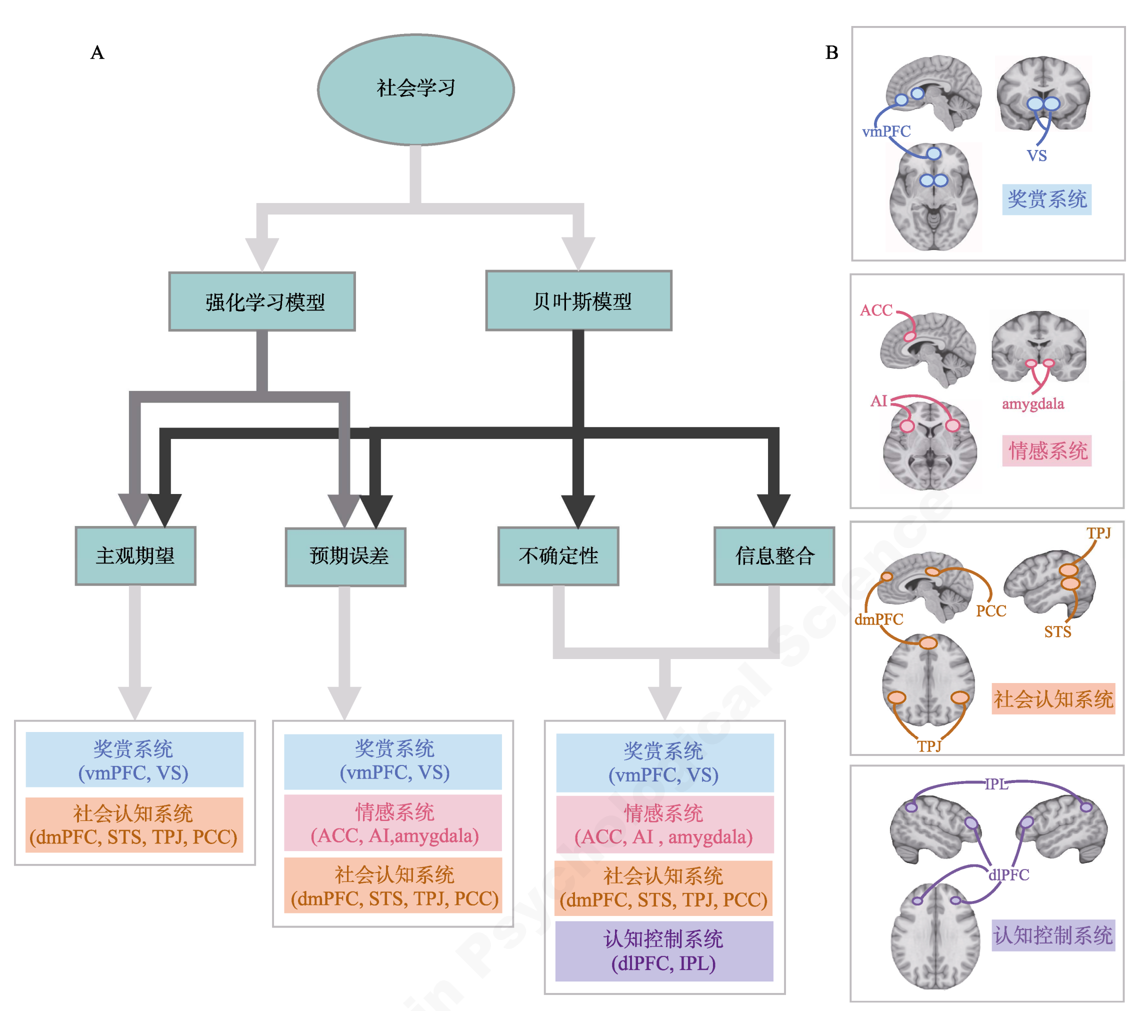


图 1　社会学习的计算神经机制

A：社会学习的计算模型及神经基础。强化学习模型下社会学习的认知过程主要包括主观期望和预期误差的计算，贝叶斯模型下社会学习的认知过程包括主观期望、预期误差的计算、不确定性的表征以及信息整合。主观期望的计算涉及奖赏系统与社会认知系统；预期误差的计算涉及奖赏系统、情感系统与社会认知系统；不确定性的表征及信息整合涉及奖赏系统、情感系统、社会认知系统与认知控制系统。B：社会学习涉及的大脑系统：奖赏系统主要包括 VS、vmPFC；情感系统主要包括 ACC、AI、amygdala；社会认知系统主要包括 dmPFC、TPJ、STS、PCC；认知控制系统主要包括 dlPFC、IPL。VS：ventral striatum，腹侧纹状体；vmPFC：ventromedial prefrontal cortex，腹内侧前额叶皮层；ACC：anterior cingulate cortex，前扣带皮层；AI：anterior insula，前侧脑岛；amygdala：杏仁核；dmPFC：dorsalmedial prefrontal cortex，背内侧前额叶皮层；TPJ：temporo-parietal junction，颞顶联合区；STS，superior temporal sulcus，颞上沟；PCC：posterior cingulate cortex，后扣带皮层；dlPFC：dorsolateral prefrontal cortex，背外侧前额叶皮层；IPL：inferior parietal lobule，顶下小叶。

知加工相关的区域。这些结果表明，即使社会学习和非社会学习共享相似的算法，执行这些算法的神经机制仍有一定的社会特异性(Lockwood, Apps, & Chang, 2020)。其次，不同的认知计算过程与大脑区域之间的关系并不是一一对应的，而是多对多的关系。例如，不同的计算过程会涉及多个脑区，而某个特定脑区会参与到多个计算过程中(图 1)。这个发现也提示，现有研究中常用的基于单变量激活的分析手段可能并不是揭示计算过程与神经活动关系的最佳技术手段。

5　展望

近年来，多个学科领域的研究者们结合实验设计、计算建模与神经影像技术，为复杂的社会学习现象提供了客观的理论模型，对相关的理论假设进行定量的公式表达，并揭示了执行相关计

算过程的大脑区域。虽然这些研究为揭示社会学习的认知计算神经机制提供了重要的证据，仍有一些非常重要的问题有待将来的研究解决。

首先，现有文献中主流的强化学习与贝叶斯模型主要是针对个体的选择(choices)进行建模，而不去考虑个体在做出某个选择时的复杂动态过程，也即忽略了个体做出决策时的反应时信息。因此，有研究者提出将学习模型与漂移扩散模型(drift diffusion model)进行结合对选择与反应时数据同时进行建模，从而更丰富地刻画个体在学习过程中的决策行为(Miletić et al., 2020; Pedersen & Frank, 2020)。漂移扩散模型假定个体在做出决定前不断在收集带有噪声的证据，直到支持某个选项的累积证据达到决策的心理阈限(Ratcliff et al., 2016)。因此，漂移扩散模型的一个重要优势是可以刻画个体决策时的动态过程，同时拟合选择数据与反应时数据。最近一项研究将漂移扩散模型扩展到社会决策情境，用来刻画个体在群体决策中对个体信息与社会信息进行整合的动态过程(Tump et al., 2020)。将来的研究可以将学习模型与漂移扩散模型相结合，以期对社会学习过程有更精细的刻画。

其次，大多数研究采用非真实的社会互动情境，可能导致较低的生态效度。例如，很多研究使用计算机程序充当被试在社会学习中的“搭档”。“程序搭档”有其独特优势：大大节约了实验成本；编程可低成本地创造具有不同特质、不同意图的被试，创造特定的社会情境，便于研究者探究自己感兴趣的内容。然而需要指出的是，社会互动的真实性是程序无法模拟的。首先，与程序互动时，被试身处于计算机前或扫描舱内，被试未必完全相信自己正在与他人进行互动。其次，尽管研究者们采用模拟算法让“程序搭档”的行为模式更加逼近于真人(Park et al., 2019)，这些算法的设计只基于研究者对社会互动的心理过程的假设，与真实社会互动情境下人们动态变化意图与策略可能相差较远。因此，这样的实验设计未必能捕捉到被试在真实社会情境下的心理过程，或损失了部分真实的心理过程。在将来的研究中，研究者应更注重实验设计的生态效度，更多采用真人互动的实验设计(Zhang & Gläscher, 2020)，以及在非真人互动的实验设计中更注重互动情境的创设。

近年来受到越来越多关注的超扫描(hyperscanning)技术或许能为解决社会学习研究中的生态效度问题提供思路。超扫描是一种通过设备连接，同步采集多个互动个体神经活动的技术，其应用范围包括功能性核磁共振技术、功能性近红外光谱技术和脑电记录等。将超扫描应用到真实、自然的多人互动中不仅可记录单一个体在社会互动中的神经机制，还可揭示社会互动中个体之间神经活动的相关性或同步性(Montague et al., 2002)，从而为社会学习的群体属性提供了保障。例如，最近有研究发现在集体任务中，群体认同感使群体内各成员前额皮层的活动同步降低，增加群体成员的盲目、冲动行为(Yang et al., 2020)。今后，研究者可以将超扫描技术普及到社会学习的研究之中，在探究单脑内神经活动的基础上，进一步探讨脑间神经活动的相互作用，从而在更具生态效度的社会情境中考察社会学习的计算神经机制，并有利于将实证研究结果推广到实践中。

再者，以往社会学习脑机制的研究大多采用单变量分析，关注单个脑区的激活与社会学习的关系，这种技术途径存在缺陷：一方面，敏感性比较弱，无法捕捉多个体素形成的联合激活模式。另一方面，传统的激活分析无法考察多个脑区间的脑网络连接模式；然而，越来越多的证据发现大脑对任务的表征与执行是通过多个脑区组成的大尺度脑网络来实现，而非局部区域的激活(Alcalá-López et al., 2018; Anderson et al., 2012; Barrett & Satpute, 2013; Bassett & Sporns, 2017; Bellucci et al., 2019; Feng et al., 2017; Park et al., 2020; van Baar et al., 2019; Wang & Olson, 2018)，但就已有文献来看，这些领域最新的神经影像数据分析手段尚未在社会学习研究领域得到良好的运用。通过对已有研究结果的回顾可知，大多对社会学习中关键心理变量计算的神经机制的发现尚停留在“脑区−计算过程”这样较单一的映射形式，从而难以发现相同脑区在不同计算过程中参与方式的差异以及一个计算过程如何调动不同脑区间的协作。总括而言，未来的研究可以更多结合多变量分析(如多变量模式分析以及表征相似性分析)与脑网络分析等技术方法从更新的角度来看待不同大脑区域在社会学习中的功能。

此外，目前关于内隐社会学习的计算神经机制研究尚少。内隐认知属于无意识的认知过程，

是复杂多面的现象，广泛存在于社会互动中。例如，许多声称反对种族歧视的人在相关的内隐联结测试中表现出明显的种族歧视倾向(Ottaway et al., 2001)。在社会学习中，人们会将和任务本身无关的结构性特征迁移到新的刺激上，但人们不会意识到这种迁移，这反映了一种内隐学习过程(Nosek et al., 2011)。例如，给被试呈现两种不同发型的女性面孔，而不同发型内隐地与吸引力的评分高低联系在一起；虽然被试没有意识到这种联系，但在测试阶段给被试呈现两种发型的新面孔时，他们会将学习阶段的评分迁移到新的面孔中(Ivanchei et al., 2019)。虽然以往有很多研究发现内隐社会学习的现象(Cone et al., 2017; Ferguson et al., 2019; Ivanchei et al., 2019; Madva & Brownstein, 2018; Panagopoulos & van der Linden, 2016; Soon, 2020)，但现有的心理学实验范式还不适合与计算建模以及脑成像技术相结合。未来的研究需要对现有的范式进行改编，从而让内隐社会学习可能涉及的计算过程能够量化。

最后，对社会学习计算与神经机制的揭示也有助于理解多种精神疾病在社会功能方面的异常。一方面，越来越多的研究发现精神疾病患者(如自闭症)在整合社会信息方面存在异常(DeMayo et al., 2019)。另一方面，计算建模与脑成像技术的结合为揭示精神疾病潜在机制提供了非常重要与新颖的方向，也即近年来发展迅速的计算精神病学(computational psychiatry)。计算精神病学在理论层面为理解精神疾病提供了量化的理论模型(Maia et al., 2017; Paulus et al., 2016)；在实践层面，相关模型拟合出的参数要比原始行为更直观地反映相应的心理过程(如冲动性)，因此可能具有更高的信度(Hedge et al., 2020)。综上，将来的研究可以结合社会学习领域中新颖的范式、计算模型以及脑成像技术来更好地理解精神疾病患者社会功能的异常，并为精神疾病的诊断以及治疗效果提供更为精细化的指标。

参考文献

高青林, 周媛. (2021). 计算模型视角下信任形成的心理和神经机制——基于信任博弈中投资者的角度. *心理科学进展, 29*(1), 178–189.

张银花, 李红, 吴寅. (2020). 计算模型在道德认知研究中的应用. *心理科学进展, 28*(7), 1042–1055.

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry, 1*, 24–57.

Alcalá-López, D., Smallwood, J., Jefferies, E., van Overwalle, F., Vogeley, K., Mars, R. B., ... Bzdok, D.. (2018). Computing the social brain connectome across systems and states. *Cerebral Cortex, 28*(7), 2207–2232.

Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology, 103*(4), 718–735.

Apps, M. A., Rushworth, M. F., & Chang, S. W. (2016). The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron, 90*(4), 692–707.

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology, 23*(3), 361–372.

Basile, B. M., Schafroth, J. L., Karaskiewicz, C. L., Chang, S. W., & Murray, E. A. (2020). The anterior cingulate cortex is necessary for forming prosocial preferences from vicarious reinforcement in monkeys. *Plos Biology, 18*(6), e3000677.

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience, 20*(3), 353–364.

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature, 456*(7219), 245–249.

Bellucci, G., Molter, F., & Park, S. Q. (2019). Neural representations of honesty predict future trust behavior. *Nature Communications, 10*(1), 1–12.

Bellucci, G., & Park, S. Q. (2020). Honesty biases trustworthiness impressions. *Journal of Experimental Psychology: General, 149*(8), 1567–1586.

Blair, K., Marsh, A. A., Morton, J., Vythilingam, M., Jones, M., Mondillo, K., ... Blair, J. R. (2006). Choosing the lesser of two evils, the better of two goods: Specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. *Journal of Neuroscience, 26*(44), 11379–11386.

Boorman, E. D., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron, 80*(6), 1558–1571.

Burke, C. J., Tobler, P. N., Schultz, W., & Baddeley, M. (2010). Striatal BOLD response reflects the impact of herd information on financial decisions. *Frontiers in Human Neuroscience, 4*, 48.

Campbell-Meiklejohn, D. K., Simonsen, A., Frith, C. D., & Daw, N. D. (2017). Independent neural computation of value from other people's confidence. *Journal of Neuroscience,*

37(3), 673–684.

Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology, 61*(2), 87–105.

Charpentier, C. J., & O'Doherty, J. P. (2018). The application of computational models to social neuroscience: Promises and pitfalls. *Social Neuroscience, 13*(6), 637–647.

Chien, S., Wiehler, A., Spezio, M., & Gläscher, J. (2016). Congruence of inherent and acquired values facilitates reward-based decision-making. *Journal of Neuroscience, 36*(18), 5003–5012.

Cohen, J. D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., ... Willke, T. L. (2017). Computational approaches to fMRI analysis. *Nature Neuroscience, 20*(3), 304–313.

Collins, A. G., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience, 21*, 576–586.

Cone, J., Mann, T. C., & Ferguson, M. J. (2017). Changing our implicit minds: How, when, and why implicit evaluations can be rapidly revised. In *Advances in experimental social psychology* (Vol. 56, pp. 131–199). Elsevier.

Corrado, G. S., Sugrue, L. P., Brown, J. R., & Newsome, W. T. (2009). The trouble with choice: Studying decision variables in the brain. In P. W. Glimcher, & E. Fehr, (Eds.), *Neuroeconomics: Decision making and the brain: Chap.30* (pp. 463-480). Londen, UK: Elsevier Academic Press.

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: A probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Computational biology, 10*(1), e1003441.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204–1215.

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience, 3*(11), 1218–1223.

de Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., & Love, B. C. (2017). Social information is integrated into value and confidence judgments according to its reliability. *Journal of Neuroscience, 37*(25), 6066–6074.

DeMayo, M. M., Young, L. J., Hickie, I. B., Song, Y. J. C., & Guastella, A. J. (2019). Circuits for social learning: A unified model and application to Autism Spectrum Disorder. *Neuroscience & Biobehavioral Reviews, 107*, 388–398.

Devaine, M., Hollard, G., & Daunizeau, J. (2014). The social Bayesian brain: Does mentalizing make a difference when we learn? *PLoS Computational Biology, 10*(12), e1003992.

Diaconescu, A. O., Mathys, C., Weber, L. A., Daunizeau, J., Kasper, L., Lomakina, E. I., ... Stephan, K. E. (2014). Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Computational Biology, 10*(9), e1003952.

Diaconescu, A. O., Mathys, C., Weber, L. A., Kasper, L., Mauer, J., & Stephan, K. E. (2017). Hierarchical prediction errors in midbrain and septum during social learning. *Social Cognitive and Affective Neuroscience, 12*(4), 618–634.

Diaconescu, A. O., Stecy, M., Kasper, L., Burke, C. J., Nagy, Z., Mathys, C., & Tobler, P. (2020). Neural Arbitration between Social and Individual Learning Systems. *eLife, 9*, e54051.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325.

FeldmanHall, O., Otto, A. R., & Phelps, E. A. (2018). Learning moral values: Another's desire to punish enhances one's own punitive behavior. *Journal of Experimental Psychology: General, 147*(8), 1211–1224.

Feng, C., Azarian, B., Ma, Y., Feng, X., Wang, L., Luo, Y. J., & Krueger, F. (2017). Mortality salience reduces the discrimination between in‐group and out‐group interactions: A functional MRI investigation using multi‐voxel pattern analysis. *Human Brain Mapping, 38*(3), 1281–1298.

Ferguson, M. J., Mann, T. C., Cone, J., & Shen, X. (2019). When and how implicit first impressions can be updated. *Current Directions in Psychological Science, 28*(4), 331–336.

Franklin, N. T., & Frank, M. J. (2015). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife, 4*, e12029.

Garvert, M. M., Moutoussis, M., Kurth-Nelson, Z., Behrens, T. E., & Dolan, R. J. (2015). Learning-induced plasticity in medial prefrontal cortex predicts preference malleability. *Neuron, 85*(2), 418–428.

Gershman, S. J. (2015). A unifying probabilistic view of associative learning. *PLoS Computational Biology, 11*(11), e1004567.

Gläscher, J. P., & O'Doherty, J. P. (2010). Model‐based approaches to neuroimaging: Combining reinforcement learning theory with fMRI data. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 501–510.

Gmytrasiewicz, P. J., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research, 24*, 49–79.

Greaves, C. J., & Farbus, L. (2006). Effects of creative and social activity on the health and well-being of socially isolated older people: Outcomes from a multi-method observational study. *The Journal of the Royal Society for the Promotion of Health, 126*(3), 134–142.

Gu, X., Wang, X., Hula, A., Wang, S., Xu, S., Lohrenz, T. M., ... Montague, P. R. (2015). Necessary, yet dissociable contributions of the insular and ventromedial prefrontal cortices to norm adaptation: Computational and lesion evidence in humans. *Journal of Neuroscience, 35*(2), 467–473.

Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: Dissociable neural correlates and effects on choice. *Nature Neuroscience, 18*(9), 1233–1235.

Hackel, L. M., & Zaki, J. (2018). Propagation of economic inequality through reciprocity and reputation. *Psychological Science, 29*(4), 604–613.

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences, 105*(18), 6741–6746.

Hedge, C., Bompas, A., & Sumner, P. (2020). Task reliability considerations in computational psychiatry. *Biological psychiatry: Cognitive Neuroscience and Neuroimaging, 5*, (9), 837–839.

Henco, L., Brandi, M.-L., Lahnakoski, J. M., Diaconescu, A. O., Mathys, C., & Schilbach, L. (2020). Bayesian modelling captures inter-individual differences in social belief computations in the putamen and insula. *Cortex, 131*, 221–236.

Hétu, S., Luo, Y., D'Ardenne, K., Lohrenz, T., & Montague, P. R. (2017). Human substantia nigra and ventral tegmental area involvement in computing social error signals during the ultimatum game. *Social Cognitive and Affective Neuroscience, 12*(12), 1972–1982.

Hill, C. A., Suzuki, S., Polania, R., Moisa, M., O'Doherty, J. P., & Ruff, C. C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature Neuroscience, 20*(8), 1142–1149.

Hill, M. R., Boorman, E. D., & Fried, I. (2016). Observational learning computations in neurons of the human anterior cingulate cortex. *Nature Communications, 7*(1), 1–12.

Hula, A., Montague, P. R., & Dayan, P. (2015). Monte carlo planning method estimates planning horizons during interactive social exchange. *PLoS Computational Biology, 11*(6), e1004254.

Hula, A., Vilares, I., Lohrenz, T., Dayan, P., & Montague, P. R. (2018). A model of risk and mental state shifts during social interaction. *PLoS Computational Biology, 14*(2), e1005935.

Ivanchei, I. I., Moroshkina, N., Tikhonov, R., & Ovchinnikova, I. (2019). Implicit learning in attractiveness evaluation: The role of conformity and analytical processing. *Journal of Experimental Psychology: General, 148*(9), 1505–1516.

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *Journal of Neuroscience, 31*(5), 1606–1613.

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. (2017). Social learning through prediction error in the brain. *NPJ Science of Learning, 2*(1), 1–9.

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., ... Casey, B. (2011). Behavioral and neural properties of social reinforcement learning. *Journal of Neuroscience, 31*(37), 13039–13045.

Khalvati, K., Mirbagheri, S., Park, S. A., Dreher, J.-C., & Rao, R. P. (2019). *A Bayesian theory of conformity in collective decision making*. Paper presented at the Advances in Neural Information Processing Systems.

Khalvati, K., Park, S. A., Mirbagheri, S., Philippe, R., Sestito, M., Dreher, J.-C., & Rao, R. P. (2019). Modeling other minds: Bayesian inference explains human choices in group decision-making. *Science Advances, 5*(11), eaax8783.

Kumar, S., Rusch, T., Doshi, P., Spezio, M., & Gläscher, J. (2019). *Modeling cooperative and competitive decision-making in the Tiger Task*. Paper presented at the The Multidisciplinary Conference on Reinforcement Learning and Decision Making.

Kumaran, D., Banino, A., Blundell, C., Hassabis, D., & Dayan, P. (2016). Computations underlying social hierarchy learning: Distinct neural mechanisms for updating and representing self-relevant information. *Neuron, 92*(5), 1135–1147.

Kuss, K., Falk, A., Trautner, P., Elger, C. E., Weber, B., & Fliessbach, K. (2013). A reward prediction error for charitable donations reveals outcome orientation of donators. *Social Cognitive and Affective Neuroscience, 8*(2), 216–223.

Lamba, A., Frank, M. J., & FeldmanHall, O. (2020). Anxiety impedes adaptive social learning under uncertainty. *Psychological Science*, *31*(5), 592–603.

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience, 20*(9), 1293–1299.

Lee, M. D., Criss, A. H., Devezer, B., Donkin, C., Etz, A., Leite, F. P., ... Vandekerckhove, J. (2019). Robust modeling in cognitive science. *Computational Brain & Behavior, 2*(3–4), 141–153.

Leong, Y. C., & Zaki, J. (2018). Unrealistic optimism in advice taking: A computational account. *Journal of Experimental Psychology: General, 147*(2), 170–189.

Li, L., Li, K. K., & Li, J. (2019). Private but not social information validity modulates social conformity bias. *Human Brain Mapping, 40*(8), 2464–2474.

Ligneul, R., Obeso, I., Ruff, C. C., & Dreher, J.–C. (2016). Dynamical representation of dominance relationships in

the human rostromedial prefrontal cortex. *Current Biology, 26*(23), 3107–3115.

Lockwood, P. L., Apps, M. A., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences, 113*(35), 9763–9768.

Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is There a 'Social' Brain? Implementations and Algorithms. *Trends in Cognitive Sciences, 24*(10), 802–813.

Lockwood, P. L., Klein-Flügge, M. C., Abdurahman, A., & Crockett, M. J. (2020). Model-free decision making is prioritized when learning to avoid harming others. *Proceedings of the National Academy of Sciences*, *117*(44), 27719-27730.

Lockwood, P. L., O'Nell, K. C., & Apps, M. A. (2020). Anterior cingulate cortex: A brain system necessary for learning to reward others? *Plos Biology, 18*(6), e3000735.

Lockwood, P. L., Wittmann, M. K., Apps, M. A., Klein-Flügge, M. C., Crockett, M. J., Humphreys, G. W., & Rushworth, M. F. (2018). Neural mechanisms for learning self and other ownership. *Nature Communications, 9*(1), 1–11.

Loughrey, D. G., Feeney, J., Kee, F., Lawlor, B. A., Woodside, J. V., Setti, A., & Power, J. M. (2020). Social factors may mediate the relationship between subjective age-related hearing loss and episodic memory. *Aging & Mental Health*, 1–8.

Madva, A., & Brownstein, M. (2018). Stereotypes, prejudice, and the taxonomy of the implicit social mind1. *Noûs, 52*(3), 611–644.

Maia, T. V., Huys, Q. J., & Frank, M. J. (2017). Theory-based computational psychiatry. *Biological Psychiatry, 82*(6), 382–384.

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience, 5*, 39.

Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience, 8*, 825.

Meshi, D., Biele, G., Korn, C. W., & Heekeren, H. R. (2012). How expert advice influences decision making. *PloS One, 7*(11), e49748.

Miletić, S., Boag, R. J., & Forstmann, B. U. (2020). Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia, 136*, 107261.

Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., ... Fisher, R. E. (2002). Hyperscanning: Simultaneous fMRI during linked social interactions. *NeuroImage*, *16*(4), 1159–1164.

Morris, R. W., Dezfouli, A., Griffiths, K. R., Le Pelley, M. E., & Balleine, B. W. (2017). The algorithmic neuroanatomy of action-outcome learning. *bioRxiv*, 137851.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences, 15*(4), 152–159.

O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annual Review of Psychology, 68*, 73–100.

O'Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452–454.

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model‐based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences, 1104*(1), 35–53.

Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitudes and racism: Effects of word familiarity and frequency on the implicit association test. *Social Cognition, 19*(2), 97–144.

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications, 6*(1), 1–14.

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences, 21*(6), 425–433.

Panagopoulos, C., & van der Linden, S. (2016). Conformity to implicit social pressure: The role of political identity. *Social Influence, 11*(3), 177–184.

Park, S. A., Goïame, S., O'Connor, D. A., & Dreher, J.-C. (2017). Integration of individual and social information for decision-making in groups of different sizes. *PLoS Biology, 15*(6), e2001958.

Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020). Map making: Constructing, combining, and inferring on abstract cognitive maps. *Neuron, 107*(6), 1–13.

Park, S. A., Sestito, M., Boorman, E. D., & Dreher, J.-C. (2019). Neural computations underlying strategic social decision-making in groups. *Nature Communications, 10*(1), 1–12.

Paulus, M. P., Huys, Q. J., & Maia, T. V. (2016). A roadmap for the development of applied computational psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(5), 386–392.

Pedersen, M. L., & Frank, M. J. (2020). Simultaneous Hierarchical Bayesian Parameter Estimation for Reinforcement Learning and Drift Diffusion Models: A Tutorial and Links to Neural Data. *Computational Brain & Behavior, 3*, 458–471.

Piray, P., Dezfouli, A., Heskes, T., Frank, M. J., & Daw, N. D.

(2019). Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Computational Biology, 15*(6), e1007043.

Piray, P., & Daw, N. D. (2020). A simple model for learning in volatile environments. *PLoS Computational Biology. 16*(7), e1007963.

Powers, A. R., Mathys, C., & Corlett, P. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science, 357*(6351), 596–600.

Pulcu, E., & Browning, M. (2019). The misestimation of uncertainty in affective disorders. *Trends in Cognitive Sciences, 23*(10), 865–875.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20*(4), 260–281.

Reiter, A. M., Suzuki, S., O'Doherty, J. P., Li, S.-C., & Eppinger, B. (2019). Risk contagion by peers affects learning and decision-making in adolescents. *Journal of Experimental Psychology: General, 148*(9), 1494–1504.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory, 2*, 64–99.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience, 15*(8), 549–562.

Rusch, T., Steixner-Kumar, S., Doshi, P., Spezio, M., & Gläscher, J. (2020). Theory of mind and decision science: Towards a typology of tasks and computational models. *Neuropsychologia*, *146*, 107488.

Seppala, E., Rossomando, T., & Doty, J. R. (2013). Social connection and compassion: Important predictors of health and well-being. *Social Research: An International Quarterly, 80*(2), 411–430.

Siegel, J. Z., Mathys, C., Rutledge, R. B., & Crockett, M. J. (2018). Beliefs about bad people are volatile. *Nature Human Behaviour, 2*(10), 750–756.

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience, 20*(10), 635–644.

Soon, V. (2020). Implicit bias and social schema: A transactive memory approach. *Philosophical Studies, 177*(7), 1857–1877.

Stanley, D. A. (2016). Getting to know you: General and specific neural computations for learning about people. *Social Cognitive and Affective Neuroscience, 11*(4), 525–536.

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2014). Absolute performance of reinforcement-learning models for the Iowa Gambling Task. *Decision, 1*(3), 161–183.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* Cambridge, MA: MIT press.

Suzuki, S., Adachi, R., Dunne, S., Bossaerts, P., & O'Doherty, J. P. (2015). Neural mechanisms underlying human consensus decision-making. *Neuron, 86*(2), 591–602.

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., ... Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron, 74*(6), 1125–1137.

Suzuki, S., Jensen, E. L., Bossaerts, P., & O'Doherty, J. P. (2016). Behavioral contagion during learning about another agent's risk-preferences acts on the neural representation of decision-risk. *Proceedings of the National Academy of Sciences, 113*(14), 3755–3760.

Suzuki, S., & O'Doherty, J. P. (2020). Breaking human social decision making into multiple components and then putting them together again. *Cortex, 127*, 221–230.

Thornton, M. A., & Tamir, D. I. (2017). Mental models accurately predict emotion transitions. *Proceedings of the National Academy of Sciences, 114*(23), 5982–5987.

Toelch, U., Bach, D. R., & Dolan, R. J. (2014). The neural underpinnings of an optimal exploitation of social information under uncertainty. *Social Cognitive and Affective Neuroscience, 9*(11), 1746–1753.

Toyokawa, W., Whalen, A., & Laland, K. N. (2019). Social learning strategies regulate the wisdom and madness of interactive crowds. *Nature Human Behaviour, 3*(2), 183–193.

Tump, A. N., Pleskac, T. J., & Kurvers, R. H. (2020). Wise or mad crowds? The cognitive mechanisms underlying information cascades. *Science Advances, 6*(29), eabb0266.

van Baar, J. M., Chang, L. J., & Sanfey, A. G. (2019). The computational and neural substrates of moral strategies in social decision-making. *Nature Communications, 10*(1), 1–14.

Wang, Y., & Olson, I. R. (2018). The original social network: White matter and social cognition. *Trends in Cognitive Sciences, 22*(6), 504–516.

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *Elife, 8*, e49547.

Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. (2016). Self-other mergence in the frontal cortex during cooperation and competition. *Neuron, 91*(2), 482–493.

Xiang, T., Lohrenz, T., & Montague, P. R. (2013). Computational substrates of norms and their violations during social exchange. *Journal of Neuroscience, 33*(3), 1099–1108.

Xiang, T., Ray, D., Lohrenz, T., Dayan, P., & Montague, P. R. (2012). Computational phenotyping of two-person interactions reveals differential neural response to depth-of-thought. *PLoS Computational Biology, 8*(12), e1002841.

Yang, J., Zhang, H., Ni, J., de Dreu, C. K., & Ma, Y. (2020).

Within-group synchronization in the prefrontal cortex associates with intergroup conflict. *Nature Neuroscience, 23*(6), 754–760.

Yoshida, W., Seymour, B., Friston, K. J., & Dolan, R. J. (2010). Neural mechanisms of belief inference during cooperative games. *Journal of Neuroscience, 30*(32), 10744–10751.

Yu, A., & Dayan, P. (2003). *Expected and unexpected uncertainty: ACh and NE in the neocortex.* Paper presented at the Advances in neural information processing systems.

Zhang, L., & Gläscher, J. (2020). A brain network supporting social influences in human decision-making. *Science Advances, 6*(34), eabb4159.

Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience, 15*(6), 695–707.

Zhu, L., Mathewson, K. E., & Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proceedings of the National Academy of Sciences, 109*(5), 1419–1424.

The computational and neural substrates underlying social learning

LI Suiqing, CHEN Xinling, ZHAI Yuzhu, ZHANG Yijie,
ZHANG Zhixing, FENG Chunliang

(*Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education; School of Psychology, South China Normal University; Center for Studies of Psychological Application, South China Normal University; Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, Guangzhou* 510631, *China*)

Abstract: Social learning refers to the belief updates of others' personal attributes and intentions as well as social norms under certain circumstances during social interactions. Due to its critical role in human decisions and social interactions, the past years have witnessed a growing body of studies that examine computational and neural basis of social learning combining computational models and human brain imaging techniques. The current literature indicates that human social learning can be well captured by reinforcement learning model and Bayesian model, based on which four computational subcomponents have been consistently identified for social learning, including subjective expectation, prediction error, uncertainty, and information integration. These computational processes have frequently engaged the involvement of brain systems associated with reward and punishment processing (e.g. ventral striatum and ventromedial prefrontal cortex), social cognition (e.g. dorsomedial prefrontal cortex and temporo-parietal junction), and cognitive control (e.g. dorsolateral prefrontal cortex). However, it should be noted that there is no one-to-one mapping between computational processes and brain regions, suggesting that multivoxel pattern analysis and brain network analysis should be utilized in future studies to reveal how different computational processes are implemented in large-scale networks according to systems neuroscience. Moreover, future studies should try to increase the ecological validity by creating real social interactions between people and by leveraging novel neuroimaging approaches (e.g. hyperscanning techniques). Finally, more efforts are needed to unravel the neural and computational signatures of implicit social learning.

Key words: social learning, computational modeling, neuroimaging, reinforcement learning model, Bayesian model